PLANTS MODIFIED WITH MINI-CHROMOSOMES

RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application No. 60/547,256 filed February 23, 2004, which is incorporated by reference herein in its entirety.

BACKGROUND OF THE INVENTION

Two general approaches are used for introduction of new genetic information ("transformation") into cells. One approach is to introduce the new genetic information as part of another DNA molecule, referred to as an "episomal vector," or "mini-chromosome", which can be maintained as an independent unit (an episome) apart from the host chromosomal DNA molecule(s). Episomal vectors contain all the necessary DNA sequence elements required for DNA replication and maintenance of the vector within the cell. Many episomal vectors are available for use in bacterial cells (for example, see Maniatis et al., "Molecular Cloning : a LaboratoryManual, "Cold Spring Harbor Laboratory, Cold Sprint Harbor, NY.. 1982.). However, only a few episomal vectors that function in higher eukaryotic cells have been developed. Higher eukaryotic episomal vectors were primarily based on naturally occurring viruses. In higher plant systems gemini viruses are double-stranded DNA viruses that replicate through a double-stranded intermediate upon which an episomal vector could be based, although the gemini virus is limited to an approximately 800 bp insert. Although an episomal plant vector based on the Cauliflower Mosaic Virus has been developed, its capacity to carry new genetic information also is limited (Brisson et al., Nature, 310 :511,1984).

The other general method of genetic transformation involves integration of introduced DNA sequences into the recipient cell's chromosomes, permitting the new information to be replicated and partitioned to the cell's progeny as a part of the natural chromosomes. The introduced DNA usually is broken and joined together in various combinations before it is integrated at random sites into the cell's chromosome (see, for example Wigler *et al.,* Cell, 11:223, 1977). Common problems with this procedure are the rearrangement of introduced DNA sequences and unpredictable levels of expression due to the location of the transgene in the genome or so called "position effect variegation" (Shingo *et al.,* Mol. Cell. Biol., 6:1787,

1986). Further, unlike episomal DNA, integrated DNA cannot normally be precisely removed. A more refined form of integrative transformation can be achieved by exploiting naturally occurring viruses that integrate into the host's chromosomes as part of their life cycle, such as retroviruses (see Chepko *et al.,* Cell, 37:1053, 1984).

One common genetic transformation method used in higher plants is based on the transfer of bacterial DNA into plant chromosomes that occurs during infection by the phytopathogenic soil bacterium *Agrobacterium* (see Nester *et al.,* Ann. Rev. Plant Phys., 35:387-413, 1984). By substituting genes of interest for the naturally transferred bacterial sequences (called T-DNA), investigators have been able to introduce new DNA into plant cells. However, even this more "refined" integrative transformation system is limited in three major ways. First, DNA sequences introduced into plant cells using the *Agrobacterium* T-DNA system are frequently rearranged (see Jones *et al.,* Mol Gen. Genet., 207:478, 1987). Second, the expression of the introduced DNA sequences varies between individual transformants (see Jones *et al.,* EMBO J., 4:2411-2418, 1985). This variability is presumably caused by rearranged sequences and the influence of surrounding sequences in the plant chromosome (*i.e.*, position effects), as well as methylation of the transgene. Finally, insertion of extra elements into the genome can disrupt the genes, promoters or other genetic elements necessary for normal plant growth and function.

Another widely used technique to genetically transform plants involves the use of micro-projectile bombardment. In this process, a nucleic acid containing the desired genetic elements to be introduced into the plant is deposited on or in small metallic particles, e.g., tungsten, platinum, or preferably gold, which are then delivered at a high velocity into the plant tissue or plant cells. However, similar problems arise as with *Agrobacterium*-mediated gene transfer, and as noted above expression of the inserted DNA can be unpredictable and insertion of extra elements into the genome can disrupt and adversely impact plant processes.

One attractive alternative to commonly used methods of transformation is the use of an artificial chromosome. Artificial chromosomes are man-made linear or circular DNA molecules constructed in part from cis-acting DNA sequence elements that provide replication and partitioning of the constructed chromosomes (see Murray *et al.,* Nature, 305:189-193, 1983). Desired elements include: (1) origin

of replication, which are the sites for initiation of DNA replication, (2) centromeres (site of kinetochore assembly and responsible for proper distribution of replicated chromosomes into daughter cells at mitosis or meiosis), and (3) if the chromosome is linear, telomeres (specialized DNA structures at the ends of linear chromosomes that function to stabilize the ends and facilitate the complete replication of the extreme termini of the DNA molecule). An additional desired element is a chromatin organizing sequence. It is well documented that centromere function is crucial for stable chromosomal inheritance in almost all eukaryotic organisms (reviewed in Nicklas 1988). The centromere accomplishes this by attaching, via centromere binding proteins, to the spindle fibers during mitosis and meiosis, thus ensuring proper gene segregation during cell divisions.

The essential chromosomal elements for construction of artificial chromosomes have been precisely characterized in lower eukaryotic species, and more recently in mouse and human. Autonomous replication sequences (ARSs) have been isolated from unicellular fungi, including *Saccharomyces cerevisiae* (brewer's yeast) and *Schizosaccharomyces pombe* (see Stinchcomb *et al.,* 1979 and Hsiao *et al.,* 1979). An ARS behaves like a origin of replication allowing DNA molecules that contain the ARS to be replicated in concert with the rest of the genome after introduction into the cell nuclei of these fungi. DNA molecules containing these sequences replicate, but in the absence of a centromere they are not partitioned into daughter cells in a controlled fashion that ensures efficient chromosome inheritance.

Artificial chromosomes have been constructed in yeast using the three cloned essential chromosomal elements (see Murray *et al.,* Nature, 305:189-193, 1983). None of the essential components identified in unicellular organisms, however, function in higher eukaryotic systems. For example, a yeast centromere sequence will not confer stable inheritance upon vectors transformed into higher eukaryotes.

In contrast to the detailed studies done in yeast, less is known about the molecular structure of functional centromeric DNA of higher eukaryotes. Ultrastructural studies indicate that higher eukaryotic kinetochores, which are specialized complexes of proteins that form on the centromere during late prophase, are large structures (mammalian kinetochore plates are approximately 0.3 μm in diameter) which possess multiple microtubule attachment sites (reviewed in Rieder,

1982). It is therefore possible that the centromeric DNA regions of these organisms will be correspondingly large, although the minimal amount of DNA necessary for centromere function may be much smaller.

While the above studies have been useful in elucidating the structure and function of centromeres, it was not known whether information derived from lower eukaryotic or mammalian higher eukaryotic organisms would be applicable to plants. There exists a need for cloned centromeres from higher eukaryotic organisms, particularly plant organisms, which would represent a first step in production of artificial chromosomes. There further exists a need for plant cells, plants, seeds and progeny containing functional, stable, and autonomous artificial chromosomes capable of carrying a large number of different genes and genetic elements.

SUMMARY OF THE INVENTION

The invention provides for adchromosomal plants, described in further detail herein, comprising a mini-chromosome, wherein said mini-chromosome preferably has a transmission efficiency during mitotic division of at least 90%, for example, at least 95%. Additionally, these adchromosomal plants may comprise a mini-chromosome having a transmission efficiency during meiotic division of, e.g., at least 80%, at least 85%, at least 90% or at least 95%.

In one embodiment, the adchromosomal plants of the invention comprise a mini-chromosome that is 1000 kilobases or less in length. In exemplary embodiments, the adchromosomal plant comprises a mini-chromosome that is 600 kilobases or less in length or 500 kilobases or less in length.

In another embodiment, the mini-chromosome of any of the preceding adchromosomal plants of the invention comprises a site for site-specific recombination.

In an embodiment, the mini-chromosome of any of the preceding adchromosomal plants of the invention comprises a centromeric nucleic acid insert derived from a crop plant centromere. In an exemplary embodiment, the centromeric nucleic acid insert is derived from genomic DNA of a plant selected from the group consisting of *Brassica, Nicotiana, Lycopersicum, Glycine* or *Zea* species. In another exemplary embodiment, the centromeric nucleic acid insert is derived from genomic

DNA of a plant selected from the group consisting of broccoli, canola, tobacco, tomato, soybean or corn.

In another embodiment, the invention provides for the mini-chromosome of any one of the preceding adchromosomal plants, further comprising a centromeric nucleic acid insert that comprises artificially synthesized repeated nucleotide sequences. These artificially synthesized repeated nucleotide sequences may be derived from natural centromere sequences, combinations or fragments of natural centromere sequences including a combination of repeats of different lengths, a combination of different sequences, a combination of both different repeat lengths and different sequences, a combination of repeats from two or more plant species, a combination of different artificially synthesized sequences or a combination of natural centromere sequence(s) and artificially synthesized sequence(s).

The invention also provides for a mini-chromosome of any of the preceding adchromosomal plants of the invention, wherein the mini-chromosome is derived from a donor clone or a centromere clone and has substitutions, deletions, insertions, duplications or arrangements of one or more nucleotides in the mini-chromosome compared to the nucleotide sequence of the donor clone or centromere clone. In one embodiment, the mini-chromosome is obtained by passage of the mini-chromosome through one or more hosts. In another embodiment, the mini-chromosome is obtained by passage of the mini-chromosome through two or more different hosts. The host may be selected from the group consisting of viruses, bacteria, yeasts, plants, prokaryotic organisms, or eukaryotic organisms.

The invention also provides for a mini-chromosome of any of the preceding adchromosomal plants of the invention, wherein the mini-chromosome comprises one or more exogenous nucleic acids. In further exemplary embodiments, the mini-chromosome comprises at least two or more, at least three or more, at least four or more, at least five or more or at least ten or more exogenous nucleic acids.

In one embodiment, at least one exogenous nucleic acid of any of the preceding mini-chromosomes of a plant is operably linked to a heterologous regulatory sequence functional in plant cells. The invention provides for exogenous nucleic acids linked to a plant regulatory sequence. The invention also provides for exogenous nucleic acids linked to a non-plant regulatory sequence, such as an inset or

yeast regulatory sequence. Exemplary regulatory sequences comprise any one of SEQ ID NOS: 4 to 23 or a functional fragment or variant thereof.

In another embodiment, the mini-chromosome of any of the preceding adchromosomal plants comprises an exogenous nucleic acid that confers herbicide resistance, insect resistance, disease resistance, or stress resistance on the plant. The invention provides for mini-chromosomes comprising an exogenous nucleic acid that confers resistance to phosphinothricin or glyphosate herbicide. The invention also provides for mini-chromosomes comprising an exogenous nucleic acid that encodes a phosphinothricin acetyltransferase, glyphosate acetyltransferase or a mutant enoylpyruvylshikimate phosphate (EPSP) synthase.

The invention also provides for the mini-chromosome of any of the preceding adchromosomal plants comprising an exogenous nucleic acid that encodes a *Bacillus thuringiensis* crystal toxin gene or *Bacillus cereus* toxin gene. The invention further provides for the mini-chromosome of any of the preceding adchromosomal plants comprising an exogenous nucleic acid that confers resistance to drought, heat, chilling, freezing, excessive moisture, ultraviolet light, ionizing radiation, toxins, pollution, mechanical stress or salt stress. The invention also provides for a mini-chromosome of any of the preceding adchromosomal plants that comprises an exogenous nucleic acid that confers resistance to a virus, bacteria, fungi or nematode.

In another embodiment, the mini-chromosome of any of the preceding adchromosomal plants comprises an exogenous nucleic acid conferring herbicide resistance, an exogenous nucleic acid conferring insect resistance, and at least one additional exogenous nucleic acid.

The invention provides for mini-chromosomes of any of the preceding adchromosomal plants comprising an exogenous nucleic acid selected from the group consisting of a nitrogen fixation gene, a plant stress-induced gene, a nutrient utilization gene, a gene that affects plant pigmentation, a gene that encodes an antisense or ribozyme molecule, a gene encoding a secretable antigen, a toxin gene, a receptor gene, a ligand gene, a seed storage gene, a hormone gene, an enzyme gene, an interleukin gene, a clotting factor gene, a cytokine gene, an antibody gene, a growth factor gene, a transcription factor gene, a transcriptional repressor gene, a

DNA-binding protein gene, a recombination gene, a DNA replication gene, a programmed cell death gene, a kinase gene, a phosphatase gene, a G protein gene, a cyclin gene, a cell cycle control gene, a gene involved in transcription, a gene involved in translation, a gene involved in RNA processing, a gene involved in RNAi, an organellar gene, a intracellular trafficking gene, an integral membrane protein gene, a transporter gene, a membrane channel protein gene, a cell wall gene, a gene involved in protein processing, a gene involved in protein modification, a gene involved in protein degradation, a gene involved in metabolism, a gene involved in biosynthesis, a gene involved in assimilation of nitrogen or other elements or nutrients, a gene involved in controlling carbon flux, gene involved in respiration, a gene involved in photosynthesis, a gene involved in light sensing, a gene involved in organogenesis, a gene involved in embryogenesis, a gene involved in differentiation, a gene involved in meiotic drive, a gene involved in self incompatibility, a gene involved in development, a gene involved in nutrient, metabolite or mineral transport, a gene involved in nutrient, metabolite or mineral storage, a calcium-binding protein gene, or a lipid-binding protein gene.

The invention also provides for a mini-chromosome of any of the preceding adchromosomal plants comprising an exogenous enzyme gene selected from the group consisting of a gene that encodes an enzyme involved in metabolizing biochemical wastes for use in bioremediation, a gene that encodes an enzyme for modifying pathways that produce secondary plant metabolites, a gene that encodes an enzyme that produces a pharmaceutical, a gene that encodes an enzyme that improves changes the nutritional content of a plant, a gene that encodes an enzyme involved in vitamin synthesis, a gene that encodes an enzyme involved in carbohydrate, polysaccharide or starch synthesis, a gene that encodes an enzyme involved in mineral accumulation or availability, a gene that encodes a phytase, a gene that encodes an enzyme involved in fatty acid, fat or oil synthesis, a gene that encodes an enzyme involved in synthesis of chemicals or plastics, a gene that encodes an enzyme involved in synthesis of a fuel and a gene that encodes an enzyme involved in synthesis of a fragrance, a gene that encodes an enzyme involved in synthesis of a flavor, a gene that encodes an enzyme involved in synthesis of a pigment or dye, a gene that encodes an enzyme involved in synthesis of a hydrocarbon, a gene that encodes an enzyme involved in synthesis of a structural or fibrous compound, a gene

that encodes an enzyme involved in synthesis of a food additive, a gene that encodes an enzyme involved in synthesis of a chemical insecticide, a gene that encodes an enzyme involved in synthesis of an insect repellent, or a gene controlling carbon flux in a plant.

In an embodiment of the invention, the mini-chromosomes of any one of the preceding adchromosomal plants comprise n copies of a repeated nucleotide sequence, wherein n is less than 1000. In other exemplary embodiments, the mini-chromosomes of the plants comprise n copies of a repeated nucleotide sequence, wherein n is at least 5, wherein n is at least 15, or wherein n is at least 50.

In another embodiment of the invention, the mini-chromosomes of any of the preceding adchromosomal plants comprise a telomere.

The invention also provides embodiments wherein the mini-chromosome of any of the preceding adchromosomal plants is circular.

In one embodiment of the invention, any of the preceding adchromosomal plants are a monocotyledon. In another embodiment of the invention, any of the preceding adchromosomal plants are a dicotyledone. The invention also provides that the adchromosomal plants of the invention are, e.g., crop plants, cereal plants, vegetable crops, field crops, fruit and vine crops, wood or fiber crops or ornamental plants. The invention also provides exemplary adchromosomal plants that are *Brassica, Nicotiana, Lycopersicum, Glycine* or *Zea* species.

Another embodiment of the invention is a part of any of the preceding adchromosomal plants. Exemplary plant parts of the invention include a pod, root, cutting, tuber, stem, stalk, fruit, berry, nut, flower, leaf, bark, wood, epidermis, vascular tissue, organ, protoplast, crown, callus culture, petiole, petal, sepal, stamen, stigma, style, bud, meristem, cambium, cortex, pith, sheath, silk or embryo. Other exemplary plant parts are a meiocyte or gamete or ovule or pollen or endosperm of any of the preceding adchromosomal plants. Other exemplary plant parts are a seed, embryo or propagule of any of the preceding adchromosomal plants.

An embodiment of the invention is a progeny of any of the preceding adchromosomal plants of the invention. These progeny of the invention may be the result of self-breeding, cross-breeding, apomyxis or clonal propagation. In exemplary embodiments, the invention also provides for progeny that comprise a mini-

chromosome that is descended from a parental mini-chromosome that contained a centromere less than 150 kilobases in length, less than 100 kilobases in length or less than 50 kilobases in length.

In another aspect, the invention provides for methods of making a mini-chromosome for use in any of the preceding adchromosomal plants of the invention. These methods comprise identifying a centromere nucleotide sequence in a genomic DNA library using a multiplicity of diverse probes, and constructing a mini-chromosome comprising the centromere nucleotide sequence. These methods may further comprise determining hybridization scores for hybridization of the multiplicity of diverse probes to genomic clones within the genomic nucleic acid library, determining a classification for genomic clones within the genomic nucleic acid library according to the hybridization scores for at least two of the diverse probes, and selecting one or more genomic clones within one or more classifications for constructing the mini-chromosome.

In exemplary embodiments, the step of determining a classification for genomic clones within the genomic nucleic acid library may utilize the hybridization scores for at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 or more different probes. A classification may comprise a pattern of high, medium or low hybridization scores to various probes.

Exemplary embodiments of probes useful in this method include a probe that hybridizes to the centromere region of a chromosome, a probe that hybridizes to satellite repeat DNA, a probe that hybridizes to retroelement DNA, a probe that hybridizes to portions of genomic DNA that are heavily methylated, a probe that hybridizes to arrays of tandem repeats in genomic DNA, a probe that hybridizes to telomere DNA or a probe that hybridizes to a pseudogene. Other exemplary probes include, a probe that hybridizes to ribosomal DNA, a probe that hybridizes to mitochondrial DNA, or a probe that hybridizes to chloroplast DNA, for which preferably a classification comprises a low hybridization score for hybridization to said probe.

Another aspect of the invention provides for methods of making any one of the preceding adchromosomal plants comprising delivering a mini-chromosome to a plant cell using a biolistic method, wherein a particle suitable for

use in biolistic method is delivered in a liquid with the mini-chromosome, and regenerating a plant from the plant cell. The liquid may further comprise a divalent ion and a di- or poly-amine. In exemplary embodiments, the liquid comprises water, $CaCl_2$, and spermidine, and the particles are gold particles. Suitable alternatives to spermidine are, e.g., spermine or other aliphatic or conjugated di- or poly-amines such as 1, 5-diaminopentane, 1, 6-diaminohexane, 1,7-diaminoheptane, 1,8-diaminooctane, histamine or related molecules.

A further aspect of the invention provides for methods of making any of the preceding adchromosomal plant comprising co-delivering to a plant cell a mini-chromosome and a nucleic acid encoding a growth inducing gene, wherein said nucleic acid is not part of the mini-chromosome, and regenerating a plant from the plant cell. The invention further provides for methods comprising co-delivering a nucleic acid encoding a growth inducing gene is not expressed or alternatively is not present in the regenerated plant. The invention also provides for methods wherein the co-delivered nucleic acid encodes a growth inducing gene expressed during regeneration. The growth inducing gene may a plant growth regulator gene, an organogenesis-promoting gene, an embryogenesis-promoting gene or regeneration-promoting gene, such as *Agrobacterium tumefaciens* isopentenyl transferase gene, *Agrobacterium rhizogenes* isopentenyl transferase gene, *Agrobacterium tumefaciens* indole-3-acetamide hydrolase (IAAH) gene or *Agrobacterium tumefaciens* tryptophan-2-monooxygenase (IAAM) gene.

Another aspect of the invention provides for methods of using any of the preceding adchromosomal plants for a food product, a pharmaceutical product or chemical product, according to which a suitable exogenous nucleic acid is expressed in adchromosomal plants or plant cells and the plant or plant cells are grown. The plant may secrete the product into its growth environment or the product may be contained within the plant, in which case the plant is harvested and desirable products are extracted.

Thus, the invention contemplates methods of using any of the preceding adchromosomal plants to produce a modified food product, for example, by growing a plant that expresses a exogenous nucleic acid that alters the nutritional content of the plant, and harvesting or processing the corn plant.

The invention also contemplates methods of using any of the preceding adchromosomal plants to produce a recombinant protein, by growing a plant comprising a mini-chromosome that comprises an exogenous nucleic acid encoding the recombinant protein. Optionally the plant is harvested and the desired recombinant protein is isolated from the plant. Exemplary recombinant proteins include pharmaceutical proteins or industrial enzymes.

The invention also contemplates methods of using any of the preceding adchromosomal plants to produce a recombinant protein, by growing a plant comprising a mini-chromosome that comprises an exogenous nucleic acid encoding an enzyme involved in synthesis of the chemical product. Optionally the plant is harvested and the desired chemical product is isolated from the plant. Exemplary chemical products include pharmaceutical products.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is an example of a mini-chromosome vector in the present invention containing 2 genes;

Figure 2 is another example of a mini-chromosome vector in the present invention containing 4 genes;

Figure 3 is a mini-chromosome from which all bacterial sequences have been removed. In this embodiment, bacterial sequence present between or among the plant-expressed genes or other mini-chromosome sequences would be excised prior to removal of the remaining bacterial sequences, by cutting the mini-chromosome with endonuclease #1, and re-ligating the structure such that the antibiotic-resistance gene #1 has been lost.

Figure 4 shows various structural configurations by which mini-chromosome elements can be oriented with respect to each other.

Figure 5 shows the alignment of *Brassica* consensus centromere satellite repeats.

Figure 6 shows the alignment of *Glycine max* (soybean) consensus centromere satellite repeats.

DETAILED DESCRIPTION OF THE INVENTION

While this invention is susceptible of embodiment in many different forms, there is shown in the drawings, and will be described herein in detail, specific embodiments thereof with the understanding that the present disclosure is to be considered as an exemplification of the principles of the invention and is not intended to limit the invention to the specific embodiments illustrated.

The invention is based on the production of modified plants, containing functional, stable, autonomous mini-chromosomes. Such mini-chromosomes have been shown herein to be meiotically transmitted to progeny.

One aspect of the invention is related to plants containing functional, stable, autonomous mini-chromosomes, preferably carrying one or more nucleic acids exogenous to the cell. Such plants carrying mini-chromosomes are contrasted to transgenic plants whose genome has been altered by chromosomal integration of an exogenous nucleic acid. Preferably, expression of the exogenous nucleic acid, either constitutively or in response to a signal which may be a challenge or a stimulus, results in an altered phenotype of the plant.

The invention provides for mini-chromosomes comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 250, 500, 1000 or more exogenous nucleic acids.

The invention contemplates that any plants, including but not limited to monocots, dicots, gymnosperm, field crops, vegetable crops, fruit and vine crops, or any specific plants named herein, may be modified by carrying autonomous mini-chromosomes as described herein. A related aspect of the invention is plant parts or plant tissues, including pollen, silk, endosperm, ovule, seed, embryo, pods, roots, cuttings, tubers, stems, stalks, fruit, berries, nuts, flowers, leaves, bark, whole plant, plant cell, plant organ, protoplast, cell culture, or any group of plant cells organized into a structural and functional unit, any cells of which carry mini-chromosomes.

A related aspect of the invention is adchromosomal plant parts or plant tissues, including pollen, silk, endosperm, ovule, seed, embryo, pods, roots, cuttings, tubers, stems, stalks, crown, callus culture, petiole, petal, sepal, stamen, stigma, style, bud, fruit, berries, nuts, flowers, leaves, bark, wood, whole plant, plant cell, plant organ, protoplast, cell culture, or any group of plant cells organized into a structural

and functional unit. In one preferred embodiment, the exogenous nucleic acid is primarily expressed in a specific location or tissue of a plant, for example, epidermis, vascular tissue, meristem, cambium, cortex, pith, leaf, sheath, flower, root or seed. Tissue-specific expression can be accomplished with, for example, localized presence of the mini-chromosome, selective maintenance of the mini-chromosome, or with promoters that drive tissue-specific expression.

Another related aspect of the invention is meiocytes, pollen, ovules, endosperm, seed, somatic embryos, apomyctic embryos, embryos derived from fertilization, vegetative propagules and progeny of the originally adchromosomal plant and of its filial generations that retain the functional, stable, autonomous mini-chromosome. Such progeny include clonally propagated plants, embryos and plant parts as well as filial progeny from self- and cross-breeding, and from apomyxis.

Preferably the mini-chromosome is transmitted to subsequent generations of viable daughter cells during mitotic cell division with a transmission efficiency of at least 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. More preferably, the mini-chromosome is transmitted to viable gametes during meiotic cell division with a transmission efficiency of at least 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% when more than one copy of the mini-chromosome is present in the gamete mother cells of the plant. Preferably, the mini-chromosome is transmitted to viable gametes during meiotic cell division with a transmission frequency of at least 20%, 30%, 40%, 45%, 46%, 47%, 48%, or 49% when one copy of the mini-chromosome is present in the gamete mother cells of the plant. For production of seeds via sexual reproduction or by apomyxis the mini-chromosome is preferably transferred into at least 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% of viable embryos when cells of the plant contain more than one copy of the mini-chromosome. For production of seeds via sexual reproduction or by apomyxis from plants with one mini-chromosome per cell, the mini-chromosome is preferably transferred into at least 20%, 30%, 40%, 45%, 46%, 47%, 48%, or 49% of viable embryos.

Preferably, a mini-chromosome that comprises an exogenous selectable trait or exogenous selectable marker can be employed to increase the frequency in subsequent generations of adchromosomal cells, tissues, gametes, embryos, endosperm, seeds, plants or progeny. More preferably, the frequency of

transmission of mini-chromosomes into viable cells, tissues, gametes, embryos, endosperm, seeds, plants or progeny can be at least 95%, 96%, 97%, 98%, 99% or 99.5% after mitosis or meiosis by applying a selection that favors the survival of adchromosomal cells, tissues, gametes, embryos, endosperm, seeds, plants or progeny over such cells, tissues, gametes, embryos, endosperm, seeds, plants or progeny lacking the mini-chromosome.

Transmission efficiency may be measured as the percentage of progeny cells or plants that carry the mini-chromosome as measured by one of several assays taught herein including detection of reporter gene fluorescence, PCR detection of a sequence that is carried by the mini-chromosome, RT-PCR detection of a gene transcript for a gene carried on the mini-chromosome, Western analysis of a protein produced by a gene carried on the mini-chromosome, Southern analysis of the DNA (either in total or a portion thereof) carried by the mini-chromosome, fluorescence in situ hybridization (FISH) or in situ localization by repressor binding, to name a few. Any assay used to detect the presence of the mini-chromosome (or a portion of the mini-chromosome) may be used to measure the efficiency of a parental cell or plant transmits the mini-chromosome to its progeny. Efficient transmission as measured by some benchmark percentage should indicate the degree to which the mini-chromosome is stable through the mitotic and meiotic cycles.

Plants of the invention may also contain chromosomally integrated exogenous nucleic acid in addition to the autonomous mini-chromosomes. The adchromosomal plants or plant parts, including plant tissues of the invention may include plants that have chromosomal integration of some portion of the mini-chromosome in some or all cells the plant. The plant, including plant tissue or plant cell is still characterized as adchromosomal despite the occurrence of some chromosomal integration. In one aspect of the invention, the autonomous mini-chromosome can be isolated from integrated exogenous nucleic acid by crossing the adchromosomal plant containing the integrated exogenous nucleic acid with plants producing some gametes lacking the integrated exogenous nucleic acid and subsequently isolating offspring of the cross, or subsequent crosses, that are adchromosomal but lack the integrated exogenous nucleic acid. This independent segregation of the mini-chromosome is one measure of the autonomous nature of the mini-chromosome.

Another aspect of the invention relates to methods for producing and isolating such adchromosomal plants containing functional, stable, autonomous mini-chromosomes.

In one embodiment, the invention contemplates improved methods for isolating native centromere sequences. In another embodiment, the invention contemplates methods for generating variants of native or artificial centromere sequences by passage through bacterial or plant or other host cells.

In a further embodiment, the invention contemplates methods for delivering the mini-chromosome into plant cells or tissues to transform the cells or tissues.

In yet another embodiment, the invention contemplates improved methods for regenerating plants, including methods for co-delivery of growth inducing genes with mini-chromosomes. The growth delivery genes include *Agrobacterium tumefaciens* or *A rhizogenes* isopentenyl transferase (IPT) genes involved in cytokinin biosynthesis, plant isopentenyl transferase (IPT) genes involved in cytokinin biosynthesis (from any plant), *Agrobacterium tumefaciens* IAAH, IAAM genes involved in auxin biosynthesis (indole-3-acetamide hydrolase and tryptophan-2-monooxygenase, respectively), *Agrobacterium rhizogenes* rolA, rolB and rolC genes involved in root formation, *Agrobacterium tumefaciens* Aux1, Aux2 genes involved in auxin biosynthesis (indole-3-acetamide hydrolase or tryptophan-2-monooxygenase genes), *Arabidopsis thaliana* leafy cotyledon genes (e.g. Lec1, Lec2) promoting embryogenesis and shoot formation (see Stone et al., Proc. Natl Acad. Sci USA 98: 11806-11811), *Arabidopsis thaliana* ESR1 gene involved in shoot formation (see Banno et al., Plant Cell 13: 2609-2618), *Arabidopsis thaliana* PGA6/WUSCHEL gene involved in embryogenesis (see Zuo et al., Plant J. 30: 349-359).

In yet a further embodiment, the invention contemplates methods for selecting modified plant cells or plant parts containing mini-chromosomes for regeneration. Such methods include assays for identifying adchromosomal plants or cells by determining that mini-chromosomes within the modified plant cell or plant are functional, stable, and autonomous. Exemplary assays for assessing mini-chromosome performance include lineage-based inheritance assays, use of chromosome loss agents to demonstrate autonomy, global mitotic mini-chromosome

inheritance assays (sectoring assays) with or without the use of agents inducing chromosomal loss, assays measuring expression levels of marker genes in the mini-chromosome over time and space in a plant, physical assays for separation of autonomous mini-chromosomes from endogenous nuclear chromosomes of plants, molecular assays demonstrating conserved mini-chromosome structure, such as PCR, Southern blots, mini-chromosome rescue, cloning and characterization of mini-chromosome sequences present in the plant, cytological assays detecting mini-chromosome presence in the cell's genome (e.g. FISH) and meiotic mini-chromosome inheritance assays, which measure the levels of mini-chromosome inheritance into a subsequent generation of plants via meiosis and gametes, embryos, endosperm or seeds.

The invention also contemplates novel methods of screening for adchromosomal plant cells that involve use of relatively low, sub-killing concentrations of selection agent (e.g. sub-killing antibiotic concentrations), and also involve use of a screenable marker (e.g., a visible marker gene) to identify clusters of modified cells carrying the screenable marker, after which these screenable cells are manipulated to homogeneity. Another aspect of the present invention is related to methods of making and compositions of non-plant promoters for expressing genes in plants.

The invention further provides isolated promoter nucleic acid sequences comprising any one of SEQ ID NOS: 4 to 23, or fragments or variants thereof that retain expression-promoting activity. Mini-chromosomes comprising non-plant promoter sequences such as these that are operably linked to plant-expressed genes (e.g., genes that confer a different phenotype on plants), are contemplated as are plants comprising such mini-chromosomes.

Another aspect is related to methods for using exonuclease to enrich for circular mini-chromosome DNA in genomic DNA preparations.

Another aspect of the invention relates to methods for using such adchromosomal plants containing a mini-chromosome for producing food products, pharmaceutical products and chemical products by appropriate expression of exogenous nucleic acid(s) contained within the mini-chromosome(s).

It has also been shown herein that mini-chromosomes containing centromeres from one plant species, when inserted into plant cells of a different species or even a different genus or family, can be stable, functional and autonomous. For example, as shown herein, a broccoli centromere (*B. oleraceae*) is functional in a canola (*B. napus*) plant. Similarly, a tomato (*Lycopersicum*) centromere is functional in a tobacco (*Nicotiana*) plant. A soybean (*G. max*) centromere is functional in a broccoli (*B. oleraceae*) and tobacco plant. Tobacco and tomato are in the same family of Solanaceae plants. Soybean is in the Leguminoseae family and broccoli is in the Brassicaceae family. Thus, another aspect of the invention is an adchromosomal plant comprising a functional, stable, autonomous mini-chromosome that contains centromere sequence derived from a different taxonomic plant species, or derived from a different taxonomic plant species, genus, family, order or class.

Yet another aspect of the invention provides novel autonomous mini-chromosomes with novel compositions and structures which are used to transform plant cells which are in turn used to generate a plant (or multiple plants). Exemplary mini-chromosomes of the invention are contemplated to be of a size 2000 kb or less in length. Other exemplary sizes of mini-chromosomes include less than or equal to, e.g., 1500 kb, 1000 kb, 900 kb, 800 kb, 700 kb, 600 kb, 500 kb, 450 kb, 400 kb, 350 kb, 300 kb, 250 kb, 200 kb, 150 kb, 100 kb, 80 kb, 60 kb, or 40 kb in length.

In a related aspect, novel centromere compositions as characterized by sequence content, size or other parameters are provided. Preferably, the minimal size of centromeric sequence is utilized in mini-chromosome construction. Exemplary sizes include a centromeric nucleic acid insert derived from a portion of plant genomic DNA, that is less than or equal to 1000 kb, 900 kb, 800 kb, 700 kb, 600 kb, 500 kb, 400 kb, 300 kb, 200 kb, 150 kb, 100 kb, 95 kb, 90 kb, 85 kb, 80 kb, 75 kb, 70 kb, 65 kb, 60 kb, 55 kb, 50 kb, 45 kb, 40 kb, 35 kb, 30 kb, 25 kb, 20 kb, 15 kb, 10 kb, 5 kb, or 2 kb in length. For example, rescued functional variant soybean centromeric sequences have been shown to be less than 30 kb in size. Another related aspect is the novel structure of the mini-chromosome, particularly structures lacking bacterial sequences, e.g, required for bacterial propagation.

In exemplary embodiments the invention contemplates mini-chromosomes or other vectors comprising a repeated nucleotide sequence derived from a *Brassica* plant and adchromosomal plants or parts containing these mini-

chromosomes. Exemplary repeated nucleotide sequences include (1) SEQ ID NO:1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO: 51 or SEQ ID NO: 52, or fragments or variants thereof, (2) combinations of any of these *Brassica* sequences or a fragment or variant thereof with another *Brassica*-derived centromeric nucleotide sequence, (3) combinations of any of these *Brassica* sequences or a fragment or variant thereof with a centromeric nucleotide sequence derived from a different plant species, and (4) combinations of any of the above with artificially synthesized centromeric nucleotide sequences.

In exemplary embodiments the invention also contemplates mini-chromosomes or other vectors comprising a repeated nucleotide sequence derived from a *Glycine max* plant and adchromosomal plants or parts containing these mini-chromosomes. Exemplary repeated nucleotide sequences include (1) SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO:26, or fragments or variants thereof, (2) combinations of any of these soybean sequences or a fragment or variant thereof with another soybean-derived centromeric nucleotide sequence, (3) combinations of any of these soybean sequences or a fragment or variant thereof with a centromeric nucleotide sequence derived from a different plant species, and (4) combinations of any of the above with artificially synthesized centromeric nucleotide sequences.

In exemplary embodiments, the invention also contemplates mini-chromosomes or other vectors comprising fragments or variants of the genomic DNA inserts of the BAC clones [identified as BB5, SB6, TB99, ZB19, or ZB113] deposited on February 23, 2005 with the American Type Culture Collection (ATCC), P.O. Box 1549 Manassas, VA 20108, USA , under Accession Nos. _____, _____, _____, _____ and , _____, respectively], or naturally occurring descendants thereof, that retain the ability to segregate during mitotic or meiotic division as described herein, as well as adchromosomal plants or parts containing these mini-chromosomes. Other exemplary embodiments include fragments or variants of the genomic DNA inserts of any of the BAC clones identified herein, or descendants thereof, and fragments or variants of the centromeric nucleic acid inserts of any of the vectors or mini-chromosomes identified herein.

In other exemplary embodiments, the invention contemplates mini-chromosomes or other vectors comprising centromeric nucleotide sequence that when hybridized to 1, 2, 3, 4, 5, 6, 7, 8 or more of the probes described in the examples

herein, under hybridization conditions described herein, e.g. low, medium or high stringency, provides relative hybridization scores as described in the examples herein. Preferably the probes for which relative hybridization scores are described herein as 5/10 or greater are used, and a hybridization signal greater than background for one or more of these probes is used to select clones. Adchromosomal plants or parts containing such mini-chromosomes are contemplated.

The advantages of the present invention include: provision of an autonomous, independent genetic linkage group for accelerating breeding; lack of disruption of host genome; multiple gene "stacking" of large numbers of genes with a potentially unlimited payload; uniformity of genetic composition exogenous DNA sequences in plant cells and plants containing autonomous mini-chromosomes; defined genetic context for predictable gene expression; higher frequency occurrence and recovery of plant cells and plants containing stably maintained exogenous DNA due to elimination of inefficient integration step; and the ability to eliminate mini-chromosomes in any tissues.

I. Composition of mini-chromosomes and mini-chromosome construction

The mini-chromosome vector of the present invention may contain a variety of elements, including (1) sequences that function as plant centromeres, (2) one or more exogenous nucleic acids, including, for example, plant-expressed genes, (3) sequences that function as an origin of replication, which may be included in the region that functions as plant centromere, (4) optionally, a bacterial plasmid backbone for propagation of the plasmid in bacteria, (5) optionally, sequences that function as plant telomeres, (6) optionally, additional "stuffer DNA" sequences that serve to separate the various components on the mini-chromosome from each other, (7) optionally "buffer" sequences such as MARs or SARs, (8) optionally marker sequences of any origin, including but not limited to plant and bacterial origin, (9) optionally, sequences that serve as recombination sites, and (10) "chromatin packaging sequences" such as cohesion and condensing binding sites.

The mini-chromosomes of the present invention may be constructed to include various components which are novel, which include, but are not limited to, the centromere comprising novel repeating centromeric sequences, and the promoters,

particularly promoters derived from non-plant species, as described in further detail below.

The mini-chromosomes of the present invention may be constructed to include various components which are novel, which include, but are not limited to, the centromere comprising novel repeating centromeric sequences, and the promoters, particularly promoters derived from non-plant species, as described in further detail below.

Novel centromere compositions

The centromere in the mini-chromosome of the present invention may comprise novel repeating centromeric sequences. An example of the mini-chromosome in the present invention is the *Brassica* BB5R4-1 mini-chromosome. The sequences set out as SEQ ID NOS:1 to 3 are relevant to the BB5R4-1 mini-chromosomes. The centromere of the BB5R4-1 mini-chromosome is 50 kb of *Brassica* centromere DNA as determined by CHEF gel analysis. To determine the sequence composition of the centromere, the mini-chromosome was randomly sheared and small fragments were cloned for sequencing, from which 11,010 bases of sequence were obtained from the centromere insert, a 0.17x coverage of the centromere. Of this sequence 9,533 bases were composed of centromere satellite repeat, the consensus of which is shown in SEQ ID NO:2. The satellite repeat was found to be 180+/-2 bp long. The remaining 1,477 bases of mini-chromosome sequence covered a unique sequence set out as SEQ ID NO:3. This sequence is considered a sampling of the centromere content of BB5R4-1.

Additional sequence analysis of another sampling of the *Brassica* centromere content of BB5R4-1 analyzing 7 contigs (1,175,176,177,180,184) that contain 118 canrep repeats from BB5R4-1 with repeat lengths of: 113x176 bp, 1x175 bp and 4x174 bp generated the consensus sequence set out in SEQ ID NO: 51. A consensus sequence was also built from 135 tandem repeats obtained from another mini-chromosome, BB280R2-3; from the largest contig (33703 kb) spanning a total of 23782 bp. The repeat lengths are: 125x176 bp, 4x182 bp, 4x175 bp and 2x177 bp and this sequence is set out as SEQ ID NO: 52. An alignment of SEQ ID NOS: 2, 51 and 52 is set out in Figure 5.

In another example, individual satellite repeats from soybean BAC clone SB12R2-3 (SEQ ID NO: 24) showed an average of 91.3% (s.d.=11.3%) identity to each other, with specific regions showing significantly higher and lower levels of variability. Comparing the satellite repeat consensus from SB12R2-3 to that obtained from randomly sampled soybean satellite sequences ChrGm1 (SEQ ID NO: 25) and ChrGm2 (SEQ ID NO: 26), see U.S. Patent Application 20030124561: *Plant centromere compositions*) identified several bases that differed significantly (χ^2 test, $P < 0.05$). The SB12MC satellite repeats showed an average length of 91.07 ± 0.40 bp, similar to the ChrGm2 91-base consensus and differing from the ChrGm1 92-base consensus. An alignment of the of consensus centromere satellite repeats is set out in Figure 6.

Exemplary embodiments of centromere nucleic acid sequences according to the present invention include fragments or variants of the genomic DNA inserts of the BAC clones [identified as BB5, SB6, TB99, ZB19, or ZB113 deposited on February 23, 2005 with the American Type Culture Collection (ATCC), P.O. Box 1549 Manassas, VA 20108, USA , under Accession Nos. _____, _____ and , _____, respectively] that retain the ability to segregate during mitotic or meiotic division as described herein. Variants of such sequences include artificially produced modifications as described herein and modifications produced via passaging through one or more bacterial, plant or other host cells as described herein.

Vectors comprising one, two, three, four, five, six, seven, eight , nine, ten, 15 or 20 or more of the elements contained in any of the exemplary vectors described in the examples below are also contemplated.

The invention specifically contemplates the alternative use of fragments or variants (mutants) of any of the nucleic acids described herein that retain the desired activity, including nucleic acids that function as centromeres, nucleic acids that function as promoters or other regulatory control sequences, or exogenous nucleic acids. Variants may have one or more additions, substitutions or deletions of nucleotides within the original nucleotide sequence. Variants include nucleic acid sequences that are at least 50%, 55%, 60, 65, 70, 75, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identical to the original nucleic acid sequence. Variants also include nucleic acid sequences that hybridize under low, medium, high or very high stringency conditions to the original nucleic acid sequence.

Similarly, the specification also contemplates the alternative use of fragments or variants of any of the polypeptides described herein.

The comparison of sequences and determination of percent identity between two nucleotide sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (1970) J. Mol. Biol. 48:444-453 algorithm which has been incorporated into the GAP program in the GCG software package (available at www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix. Preferably parameters are set so as to maximize the percent identity.

As used herein, the term "hybridizes under low stringency, medium stringency, and high stringency conditions" describes conditions for hybridization and washing. Guidance for performing hybridization reactions can be found in Current Protocols in Molecular Biology (1989) John Wiley & Sons, N.Y., 6.3.1-6.3.6, which is incorporated by reference. Aqueous and non-aqueous methods are described in that reference and either can be used. Specific hybridization conditions referred to herein are as follows: 1) low stringency hybridization conditions in 6 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by two washes in 0.5 x SSC, 0.1% SDS, at least at 50°C; 2) medium stringency hybridization conditions in 6 x SSC at about 45°C, followed by one or more washes in 0.2 x SSC, 0.1% SDS at 55°C; 3) high stringency hybridization conditions in 6 x SSC at about 45°C, followed by one or more washes in 0.2 x SSC, 0.1% SDS at 65°C.

Mini-chromosome sequence content and structure

Plant-expressed genes from non-plant sources may be modified to accommodate plant codon usage, to insert preferred motifs near the translation initiation ATG codon, to remove sequences recognized in plants as 5' or 3' splice sites, or to better reflect plant GC/AT content. Plant genes typically have a GC content of more than 35%, and coding sequences which are rich in A and T nucleotides can be problematic. For example, ATTTA motifs may destabilize mRNA; plant polyadenylation signals such as AATAAA at inappropriate positions within the message may cause premature truncation of transcription; and monocotyledons may recognize AT-rich sequences as splice sites.

Each exogenous nucleic acid or plant-expressed gene may include a promoter, a coding region and a terminator sequence, which may be separated from each other by restriction endonuclease sites or recombination sites or both. Genes may also include introns, which may be present in any number and at any position within the transcribed portion of the gene, including the 5' untranslated sequence, the coding region and the 3' untranslated sequence. Introns may be natural plant introns derived from any plant, or artificial introns based on the splice site consensus that has been defined for plant species. Some intron sequences have been shown to enhance expression in plants. Optionally the exogenous nucleic acid may include a plant transcriptional terminator, non-translated leader sequences derived from viruses that enhance expression, a minimal promoter, or a signal sequence controlling the targeting of gene products to plant compartments or organelles.

The coding regions of the genes can encode any protein, including but not limited to visible marker genes (for example, fluorescent protein genes, other genes conferring a visible phenotype to the plant) or other screenable or selectable marker genes (for example, conferring resistance to antibiotics, herbicides or other toxic compounds or encoding a protein that confers a growth advantage to the cell expressing the protein) or genes which confer some commercial or agronomic value to the adchromosomal plant. Multiple genes can be placed on the same mini-chromosome vector, limited only by the number of restriction endonuclease sites or site-specific recombination sites present in the vector. The genes may be separated from each other by restriction endonuclease sites, homing endonuclease sites, recombination sites or any combinations thereof. Any number of genes can be present. Figures 1 and 2 show mini-chromosome vector structures with 2 and 4 genes, respectively.

The mini-chromosome vector may also contain a bacterial plasmid backbone for propagation of the plasmid in bacteria such as *E. coli, A. tumefaciens*, or *A. rhizogenes*. The plasmid backbone may be that of a low-copy vector or in other embodiments it may be desirable to use a mid to high level copy backbone. In one embodiment of the invention, this backbone contains the replicon of the F' plasmid of *E. coli*. However, other plasmid replicons, such as the bacteriophage P1 replicon, or other low-copy plasmid systems such as the RK2 replication origin, may also be used. The backbone may include one or several antibiotic-resistance genes conferring

resistance to a specific antibiotic to the bacterial cell in which the plasmid is present. Bacterial antibiotic-resistance genes include but are not limited to kanamycin-, ampicillin-, chloramphenicol-, streptomycin-, spectinomycin-, tetracycline- and gentamycin-resistance genes.

The mini-chromosome vector may also contain plant telomeres. An exemplary telomere sequence is TTTAGGG or its complement. Telomeres are specialized DNA structures at the ends of linear chromosomes that function to stabilize the ends and facilitate the complete replication of the extreme termini of the DNA molecule (Richards et. al., Cell.1988 Apr 8;53(1):127-36; Ausubel et al., Current Protocols in Molecular Biology, Wiley & Sons, 1997).

Additionally, the mini-chromosome vector may contain "stuffer DNA" sequences that serve to separate the various components on the mini-chromosome (centromere, genes, telomeres) from each other. The stuffer DNA may be of any origin, prokaryotic or eukaryotic, and from any genome or species, plant, animal, microbe or organelle, or may be of synthetic origin. The stuffer DNA can range from 100 bp to 10 Mb in length and can be repetitive in sequence, with unit repeats from 10 to 1,000,000 bp. Examples of repetitive sequences that can be used as stuffer DNAs include but are not limited to: rDNA, satellite repeats, retroelements, transposons, pseudogenes, transcribed genes, microsatellites, tDNA genes, short sequence repeats and combinations thereof. Alternatively, the stuffer DNA can consist of unique, non-repetitive DNA of any origin or sequence. The stuffer sequences may also include DNA with the ability to form boundary domains, such as but not limited to scaffold attachment regions (SARs) or matrix attachment regions (MARs). The stuffer DNA may be entirely synthetic, composed of random sequence. In this case, the stuffer DNA may have any base composition, or any A/T or G/C content. For example, the G/C content of the stuffer DNA could resemble that of the plant (~30-40%), or could be much lower (0-30%) or much higher (40-100%). Alternatively, the stuffer sequences could be synthesized to contain an excess of any given nucleotide such as A, C, G or T. Different synthetic stuffers of different compositions may also be combined with each other. For example a fragment with low G/C content may be flanked or abutted by a fragment of medium or high G/C content, or vice versa.

In one embodiment of the invention, the mini-chromosome has a circular structure without telomeres, as shown in Figures 1 and 2 "circular". In

another embodiment, the mini-chromosome has a circular structure with telomeres, as shown in Figures 1 and 2 "linear". In a third embodiment, the mini-chromosome has a linear structure with telomeres, as would result if the "linear" structure shown in Figures 1 and 2 were to be cut with a unique endonuclease, exposing the telomeres at the ends of a DNA molecule that contains all of the sequence contained in the original, closed construct with the exception of the antibiotic-resistance gene #1. In a fourth embodiment of the invention, the telomeres could be placed in such a manner that the bacterial replicon, backbone sequences, antibiotic-resistance genes and any other sequences of bacterial origin and present for the purposes of propagation of the mini-chromosome in bacteria, can be removed from the plant-expressed genes, the centromere, telomeres, and other sequences by cutting the structure with unique endonuclease #2 (Figure 3). This results in a mini-chromosome from which much of, or preferably all, bacterial sequences have been removed. In this embodiment, bacterial sequence present between or among the plant-expressed genes or other mini-chromosome sequences would be excised prior to removal of the remaining bacterial sequences by cutting the mini-chromosome with homing endonuclease #1, and re-ligating the structure such that the antibiotic-resistance gene #1 has been lost (Figure 3). In all of the structures shown in figures 1, 2 and 3, the unique endonuclease site may be the recognition sequence of a homing endonuclease. Alternatively, the endonucleases and their sites can be replaced with any specific DNA cutting mechanism and its specific recognition site such as rare-cutting endonuclease or recombinase and its specific recognition site, as long as that site is present in the mini-chromosomes only at the indicated positions.

Various structural configurations are possible by which mini-chromosome elements can be oriented with respect to each other. A centromere can be placed on a mini-chromosome either between genes or outside a cluster of genes next to one telomere or next to the other telomere. Stuffer DNAs can be combined with these configurations to place the stuffer sequences inside the telomeres, around the centromere between genes or any combination thereof. Thus, a large number of alternative mini-chromosome structures are possible, depending on the relative placement of centromere DNA, genes, stuffer DNAs, bacterial sequences, telomeres, and other sequences. The sequence content of each of these variants is the same, but their structure may be different depending on how the sequences are placed. These

variations in architecture are possible both for linear and for circular mini-chromosomes.

Exemplary centromere components

Centromere components may be isolated or derived from native plant genome, for example, modified through recombinant techniques or through the cell-based techniques described below. Alternatively, wholly artificial centromere components may be constructed using as a general guide the sequence of native centromeres. Combinations of centromere components derived from natural sources and/or combinations of naturally derived and artificial components are also contemplated As noted above, centromere sequence from one taxonomic plant species has been shown to be functional in another taxonomic plant species, genus and family.

In one embodiment, the centromere contains n copies of a repeated nucleotide sequence obtained by the methods disclosed herein, wherein n is at least 2. In another embodiment, the centromere contains n copies of interdigitated repeats. An interdigitated repeat is a DNA sequence that consists of two distinct repetitive elements that combine to create an unique permutation. Potentially any number of repeat copies capable of physically being placed on the recombinant construct could be included on the construct, including about 5, 10, 15, 20, 30, 50, 75, 100, 150, 200, 300, 400, 500, 750, 1,000, 1,500, 2,000, 3,000, 5,000, 7,500, 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000 and about 100,000, including all ranges in-between such copy numbers. Moreover, the copies, while largely identical, can vary from each other. Such repeat variation is commonly observed in naturally occurring centromeres. The length of the repeat may vary, but will preferably range from about 20 bp to about 360 bp, from about 20 bp to about 250 bp, from about 50 bp to about 225 bp, from about 75 bp to about 210 bp, such as a 92 bp repeat and a 97 bp repeat, from about 100 bp to about 205 bp, from about 125 bp to about 200 bp, from about 150 bp to about 195 bp, from about 160 bp to about 190 and from about 170 bp to about 185 bp including about 180 bp.

Modification of centromeres isolated from native plant genome

Modification and changes may be made in the centromeric DNA segments of the current invention and still obtain a functional molecule with desirable

characteristics. The following is a discussion based upon changing the nucleic acids of a centromere to create an equivalent, or even an improved, second generation molecule.

In particular embodiments of the invention, mutated centromeric sequences are contemplated to be useful for increasing the utility of the centromere. It is specifically contemplated that the function of the centromeres of the current invention may be based in part of in whole upon the secondary structure of the DNA sequences of the centromere, modification of the DNA with methyl groups or other adducts, and / or the proteins which interact with the centromere. By changing the DNA sequence of the centromere, one may alter the affinity of one or more centromere-associated protein(s) for the centromere and / or the secondary structure or modification of the centromeric sequences, thereby changing the activity of the centromere. Alternatively, changes may be made in the centromeres of the invention which do not affect the activity of the centromere. Changes in the centromeric sequences which reduce the size of the DNA segment needed to confer centromere activity are contemplated to be particularly useful in the current invention, as would changes which increased the fidelity with which the centromere was transmitted during mitosis and meiosis.

Modification of centromeres by passage through bacteria, plant or other hosts or processes

In the methods of the present invention, the resulting mini-chromosome DNA sequence may also be a derivative of the parental clone or centromere clone having substitutions, deletions, insertions, duplications and/or rearrangements of one or more nucleotides in the nucleic acid sequence. Such nucleotide mutations may occur individually or consecutively in stretches of 1, 2, 3, 4, 5, 10, 20, 40, 80, 100, 200, 400, 800, 1000, 2000, 4000, 8000, 10000, 50000, 100000, and about 200000, including all ranges in-between.

Variations of mini-chromosomes may arise through passage of mini-chromosomes through various hosts including virus, bacteria, yeast, plant or other prokaryotic or eukaryotic organism and may occur through passage of multiple hosts or individual host. Variations may also occur by replicating the mini-chromosome in vitro.

Derivatives may be identified through sequence analysis, or variations in mini-chromosome molecular weight through electrophoresis such as, but not limited to, CHEF gel analysis, column or gradient separation, or any other methods used in the field to determine and/or analyze DNA molecular weight or sequence content. Alternately, derivatives may be identified by the altered activity of a derivative in conferring centromere function to a mini-chromosome.

Exemplary exogenous nucleic acids including plant-expressed genes

Of particular interest in the present invention are exogenous nucleic acids which when introduced into plants alter the phenotype of the plant, a plant organ, plant tissue, or portion of the plant. Exemplary exogenous nucleic acids encode polypeptides involved in one or more important biological properties in plants. Other exemplary exogenous nucleic acids alter expression of exogenous or endogenous genes, either increasing or decreasing expression, optionally in response to a specific signal or stimulus.

As used herein, the term "trait" can refer either to the altered phenotype of interest or the nucleic acid which causes the altered phenotype of interest.

One of the major purposes of transformation of crop plants is to add some commercially desirable, agronomically important traits to the plant. Such traits include, but are not limited to, herbicide resistance or tolerance; insect (pest) resistance or tolerance; disease resistance or tolerance (viral, bacterial, fungal, nematode or other pathogens); stress tolerance and/or resistance, as exemplified by resistance or tolerance to drought, heat, chilling, freezing, excessive moisture, salt stress, mechanical stress, extreme acidity, alkalinity, toxins, UV light, ionizing radiation or oxidative stress; increased yields, whether in quantity or quality; enhanced or altered nutrient acquisition and enhanced or altered metabolic efficiency; enhanced or altered nutritional content and makeup of plant tissues used for food, feed, fiber or processing; physical appearance; male sterility; drydown; standability; prolificacy; starch quantity and quality; oil quantity and quality; protein quality and quantity; amino acid composition; modified chemical production; altered pharmaceutical or nutraceutical properties; altered bioremediation properties; increased biomass; altered growth rate; altered fitness; altered biodegradability; altered CO_2 fixation; presence of bioindicator activity; altered digestibility by humans

or animals; altered allergenicity; altered mating characteristics; altered pollen dispersal; improved environmental impact; altered nitrogen fixation capability; the production of a pharmaceutically active protein; the production of a small molecule with medicinal properties; the production of a chemical including those with industrial utility; the production of nutraceuticals, food additives, carbohydrates, RNAs, lipids, fuels, dyes, pigments, vitamins, scents, flavors, vaccines, antibodies, hormones, and the like; and alterations in plant architecture or development, including changes in developmental timing, photosynthesis, signal transduction, cell growth, reproduction, or differentiation. Additionally one could create a library of an entire genome from any organism or organelle including mammals, plants, microbes, fungi, or bacteria, represented on mini-chromosomes.

In one embodiment, the modified plant may exhibit increased or decreased expression or accumulation of a product of the plant, which may be a natural product of the plant or a new or altered product of the plant. Exemplary products include an enzyme, an RNA molecule, a nutritional protein, a structural protein, an amino acid, a lipid, a fatty acid, a polysaccharide, a sugar, an alcohol, an alkaloid, a carotenoid, a propanoid, a phenylpropanoid, or terpenoid, a steroid, a flavonoid, a phenolic compound, an anthocyanin, a pigment, a vitamin or a plant hormone. In another embodiment, the modified plant has enhanced or diminished requirement for light, water, nitrogen, or trace elements. In another embodiment the modified plant has an enhance ability to capture or fix nitrogen from its environment. In yet another embodiment, the modified plant is enriched for an essential amino acid as a proportion of a protein fraction of the plant. The protein fraction may be, for example, total seed protein, soluble protein, insoluble protein, water-extractable protein, and lipid-associated protein. The modification may include overexpression, underexpression, antisense modulation, sense suppression, inducible expression, inducible repression, or inducible modulation of a gene.

A brief summary of exemplary improved properties and polypeptides of interest for either increased or decreased expression is provided below.

(i) Herbicide Resistance

A herbicide resistance (or tolerance) trait is a characteristic of a modified plant that is resistant to dosages of an herbicide that is typically lethal to a